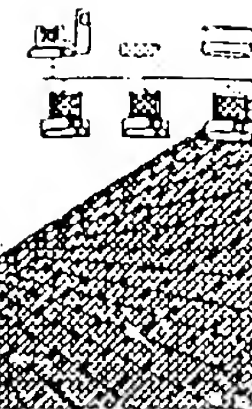# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: *10/656 531*

Source *IFWO*

Date Processed by STIC: *3/22/04*

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT
MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221
Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 4.1 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS.

http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses.

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1D03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 10|656,531

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 ____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use **space characters**, instead.

4 ____ Non-ASCII

The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ____ Variable Length

Sequence(s) ____ contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

7 ____ Skipped Sequences (OLD RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.

8 ____ Skipped Sequences (NEW RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of **n** or **Xaa**, and which residue n or **Xaa** represents.

10 ____ Invalid <213> Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 ____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n/Xaa

"n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid

CRFI

IFWO

RAW SEQUENCE LISTING   DATE: 03/22/2004
PATENT APPLICATION: **US/10/656,531**   TIME: 07:03:47

Input Set : **N:\efs\10656531\CTCHP01016-usbios-S000001.txt**
Output Set: **N:\CRF4\03222004\J656531.raw**

```
 3 <110> APPLICANT: Baltimore et al.
 5 <120> TITLE OF INVENTION: CHIMERIC NUCLEASES TO STIMULATE GENE TARGETING
 7 <130> FILE REFERENCE: CTCH-P01-016
 9 <140> CURRENT APPLICATION NUMBER: 10/656,531
10 <141> CURRENT FILING DATE: 2003-09-05
12 <150> PRIOR APPLICATION NUMBER: 60/408,454
13 <151> PRIOR FILING DATE: 2002-09-05
15 <150> PRIOR APPLICATION NUMBER: 60/419,341
16 <151> PRIOR FILING DATE: 2002-10-17
18 <150> PRIOR APPLICATION NUMBER: 60/484,788
19 <151> PRIOR FILING DATE: 2003-07-03
21 <160> NUMBER OF SEQ ID NOS: 52
23 <170> SOFTWARE: PatentIn version 3.2
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 18
27 <212> TYPE: DNA
28 <213> ORGANISM: primer
30 <400> SEQUENCE: 1
31 tagggataac agggtaat                                        18
34 <210> SEQ ID NO: 2
35 <211> LENGTH: 9
36 <212> TYPE: DNA
37 <213> ORGANISM: primer
39 <400> SEQUENCE: 2
40 ggggaagaa                                                  9
43 <210> SEQ ID NO: 3
44 <211> LENGTH: 9
45 <212> TYPE: DNA
46 <213> ORGANISM: primer
48 <400> SEQUENCE: 3
49 gcgtggtcg                                                  9
52 <210> SEQ ID NO: 4
53 <211> LENGTH: 9
54 <212> TYPE: DNA
55 <213> ORGANISM: primer
57 <400> SEQUENCE: 4
58 gatgttgct                                                  9
61 <210> SEQ ID NO: 5
62 <211> LENGTH: 15
63 <212> TYPE: DNA
64 <213> ORGANISM: primer
67 <220> FEATURE:
68 <221> NAME/KEY: misc_feature
```

[handwritten annotations:]

Does Not Comply
Corrected Diskette Needed
(pg. 1-5)

primer — Invalid Response, please see item #10 on error summary sheet.

primer — Invalid Response, please see item #10 on error summary sheet.

primer — Invalid Response

primer — Invalid Response

primer — Invalid Response

RAW SEQUENCE LISTING   DATE: 03/22/2004
PATENT APPLICATION: US/10/656,531   TIME: 07:03:47

Input Set : N:\efs\10656531\CTCHP01016-usbios-S000001.txt
Output Set: N:\CRF4\03222004\J656531.raw

```
  69 <222> LOCATION: (5)..(10)
  70 <223> OTHER INFORMATION: n is a, c, g, or t
  72 <400> SEQUENCE: 5
W--> 73 gatgnnnnnn ttgct                                          15
  76 <210> SEQ ID NO: 6
  77 <211> LENGTH: 9
  78 <212> TYPE: DNA
  79 <213> ORGANISM: primer          — SAme error
  82 <220> FEATURE:
  83 <221> NAME/KEY: misc_feature
  84 <222> LOCATION: (2)..(3)
  85 <223> OTHER INFORMATION: n is a, c, g, or t
  87 <220> FEATURE:
  88 <221> NAME/KEY: misc_feature
  89 <222> LOCATION: (5)..(6)
  90 <223> OTHER INFORMATION: n is a, c, g, or t
  92 <220> FEATURE:
  93 <221> NAME/KEY: misc_feature
  94 <222> LOCATION: (8)..(9)
  95 <223> OTHER INFORMATION: n is a, c, g, or t
  97 <400> SEQUENCE: 6
W--> 98 gnngnngnn                                                  9
 101 <210> SEQ ID NO: 7
 102 <211> LENGTH: 24
 103 <212> TYPE: DNA
 104 <213> ORGANISM: primer          ~ SAme error
 107 <220> FEATURE:
 108 <221> NAME/KEY: misc_feature
 109 <222> LOCATION: (1)..(2)
 110 <223> OTHER INFORMATION: n is a, c, g, or t
 112 <220> FEATURE:
 113 <221> NAME/KEY: misc_feature
 114 <222> LOCATION: (4)..(5)
 115 <223> OTHER INFORMATION: n is a, c, g, or t
 117 <220> FEATURE:
 118 <221> NAME/KEY: misc_feature
 119 <222> LOCATION: (7)..(8)
 120 <223> OTHER INFORMATION: n is a, c, g, or t
 122 <220> FEATURE:
 123 <221> NAME/KEY: misc_feature
 124 <222> LOCATION: (10)..(15)
 125 <223> OTHER INFORMATION: n is a, c, g, or t
 127 <220> FEATURE:
 128 <221> NAME/KEY: misc_feature
 129 <222> LOCATION: (17)..(18)
 130 <223> OTHER INFORMATION: n is a, c, g, or t
 132 <220> FEATURE:
 133 <221> NAME/KEY: misc_feature
 134 <222> LOCATION: (20)..(21)
```

RAW SEQUENCE LISTING   DATE: 03/22/2004
PATENT APPLICATION: US/10/656,531   TIME: 07:03:47

Input Set : **N:\efs\10656531\CTCHP01016-usbios-S000001.txt**
Output Set: **N:\CRF4\03222004\J656531.raw**

```
135 <223> OTHER INFORMATION: n is a, c, g, or t
137 <220> FEATURE:
138 <221> NAME/KEY: misc_feature
139 <222> LOCATION: (23)..(24)
140 <223> OTHER INFORMATION: n is a, c, g, or t
142 <400> SEQUENCE: 7
```
W--> **143** nncnncnncn nnnnngnngn ngnn   **24**

same errors

```
146 <210> SEQ ID NO: 8
147 <211> LENGTH: 9
148 <212> TYPE: DNA
149 <213> ORGANISM: primer
151 <400> SEQUENCE: 8
152 ggggaagaa                                                            9
155 <210> SEQ ID NO: 9
156 <211> LENGTH: 8
157 <212> TYPE: DNA
158 <213> ORGANISM: primer
160 <400> SEQUENCE: 9
161 gcgtggcg                                                             8
164 <210> SEQ ID NO: 10
165 <211> LENGTH: 24
166 <212> TYPE: DNA
167 <213> ORGANISM: primer
169 <400> SEQUENCE: 10
170 accatcttct tcaaggacga cggc                                          24
173 <210> SEQ ID NO: 11
174 <211> LENGTH: 9
175 <212> TYPE: DNA
176 <213> ORGANISM: primer
178 <400> SEQUENCE: 11
179 gaagatggt                                                            9
182 <210> SEQ ID NO: 12
183 <211> LENGTH: 9
184 <212> TYPE: DNA
185 <213> ORGANISM: primer
187 <400> SEQUENCE: 12
188 gacgacggc                                                            9
191 <210> SEQ ID NO: 13
192 <211> LENGTH: 24
193 <212> TYPE: DNA
194 <213> ORGANISM: primer
196 <400> SEQUENCE: 13
197 ggcgcccacc atcgcgtcgc agcc                                          24
200 <210> SEQ ID NO: 14
201 <211> LENGTH: 9
202 <212> TYPE: DNA
203 <213> ORGANISM: primer
205 <400> SEQUENCE: 14
206 gtgggcgcc                                                            9
```

RAW SEQUENCE LISTING   DATE: 03/22/2004
PATENT APPLICATION: **US/10/656,531**   TIME: 07:03:47

Input Set : **N:\efs\10656531\CTCHP01016-usbios-S000001.txt**
Output Set: **N:\CRF4\03222004\J656531.raw**

```
209 <210> SEQ ID NO: 15
210 <211> LENGTH: 9
211 <212> TYPE: DNA
212 <213> ORGANISM: primer
214 <400> SEQUENCE: 15
215 gtcgcagcc                                                        9
218 <210> SEQ ID NO: 16
219 <211> LENGTH: 168
220 <212> TYPE: DNA
221 <213> ORGANISM: primer
223 <400> SEQUENCE: 16
224 actagcaacc tcaaacagac accatggtgc atctgactcc tgaggagaag tctgccgtta    60
226 ctgccctgtg gggcaaggtg aacgtgatcg ttggagtttg tctgtggtac cacgtagact   120
228 gaggactcct cttcagacgg caatgacggg acaccccgtt ccacttgc               168
231 <210> SEQ ID NO: 17
232 <211> LENGTH: 9
233 <212> TYPE: DNA
234 <213> ORGANISM: primer
236 <400> SEQUENCE: 17
237 gaggttgct                                                         9
240 <210> SEQ ID NO: 18
241 <211> LENGTH: 7
242 <212> TYPE: PRT
243 <213> ORGANISM: peptide
245 <400> SEQUENCE: 18
247 Gln Ser Ser Asp Leu Thr Arg
248 1               5
251 <210> SEQ ID NO: 19
252 <211> LENGTH: 7
253 <212> TYPE: PRT
254 <213> ORGANISM: peptide
256 <400> SEQUENCE: 19
258 Thr Ser Gly Glu Leu Val Arg
259 1               5
262 <210> SEQ ID NO: 20
263 <211> LENGTH: 7
264 <212> TYPE: PRT
265 <213> ORGANISM: peptide
267 <400> SEQUENCE: 20
269 Thr Ser Gly Ala Leu Thr Arg
270 1               5
273 <210> SEQ ID NO: 21
274 <211> LENGTH: 7
275 <212> TYPE: PRT
276 <213> ORGANISM: peptide
278 <400> SEQUENCE: 21
280 Thr Ser Gly Ser Leu Thr Arg
281 1               5
284 <210> SEQ ID NO: 22
```

Invalid Response, please see item #10 on error summary sheet.

Please see item #10 on error summary sheet.

Same errors

RAW SEQUENCE LISTING   DATE: 03/22/2004
PATENT APPLICATION: US/10/656,531   TIME: 07:03:47

Input Set : N:\efs\10656531\CTCHP01016-usbios-S000001.txt
Output Set: N:\CRF4\03222004\J656531.raw

same errors

```
285 <211> LENGTH: 7
286 <212> TYPE: PRT
287 <213> ORGANISM: peptide
289 <400> SEQUENCE: 22
291 Arg Ser Asp Asn Leu Thr Arg
292 1               5
295 <210> SEQ ID NO: 23
296 <211> LENGTH: 7
297 <212> TYPE: PRT
298 <213> ORGANISM: peptide
300 <400> SEQUENCE: 23
302 Arg Ser Asp Asn Leu Val Arg
303 1               5
306 <210> SEQ ID NO: 24
307 <211> LENGTH: 48
308 <212> TYPE: DNA
309 <213> ORGANISM: primer
311 <400> SEQUENCE: 24
312 agcaacctct ctagagcgtg ggcgtcgttg gagagatctc gcacccgc          48
315 <210> SEQ ID NO: 25
316 <211> LENGTH: 9
317 <212> TYPE: DNA
318 <213> ORGANISM: primer
320 <400> SEQUENCE: 25
321 ggcaaggtg                                                      9
324 <210> SEQ ID NO: 26
325 <211> LENGTH: 7
326 <212> TYPE: PRT
327 <213> ORGANISM: peptide
329 <400> SEQUENCE: 26
331 Arg Ser Asp Ala Leu Thr Arg
332 1               5
335 <210> SEQ ID NO: 27
336 <211> LENGTH: 7
337 <212> TYPE: PRT
338 <213> ORGANISM: peptide
340 <400> SEQUENCE: 27
342 Arg Ser Asn Ser Leu Thr Arg
343 1               5
346 <210> SEQ ID NO: 28
347 <211> LENGTH: 7
348 <212> TYPE: PRT
349 <213> ORGANISM: peptide
351 <400> SEQUENCE: 28
353 Arg Ser Asp Thr Leu Ser Asn
354 1               5
357 <210> SEQ ID NO: 29
358 <211> LENGTH: 7
359 <212> TYPE: PRT
```

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 03/22/2004
PATENT APPLICATION: US/10/656,531   TIME: 07:03:48

Input Set : **N:\efs\10656531\CTCHP01016-usbios-S000001.txt**
Output Set: **N:\CRF4\03222004\J656531.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos: 5,6,7,8,9,10
Seq#:6; N Pos. 2,3,5,6,8,9
Seq#:7; N Pos. 1,2,4,5,7,8,10,11,12,13,14,15,17,18,20,21,23,24

VERIFICATION SUMMARY　　　　　　　　　DATE: 03/22/2004
PATENT APPLICATION: **US/10/656,531**　　　TIME: 07:03:48

Input Set : **N:\efs\10656531\CTCHP01016-usbios-S000001.txt**
Output Set: **N:\CRF4\03222004\J656531.raw**

L:73 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:98 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0